

# SYSTEMS AND METHODS FOR REDUCING BOUNDARY ARTIFACTS IN HYBRID COMPRESSION

## BACKGROUND OF THE INVENTION

### 1. Field of Invention

This invention generally relates to processing compressed digital images. More particularly, this invention relates to methods and systems which reduce the boundary ringing artifacts occurring in hybrid compression.

### 2. Description of Related Art

Data compression is required in data handling processes, where too much data is present for practical applications using the data. Commonly, compression is used in communication links to reduce the transmission time or required bandwidth. Similarly, compression is preferred in image storage systems, including digital printers and copiers, where "pages" of a document to be printed are stored temporarily in precollation memory. The amount of media space on which the image data is stored can be substantially reduced with compression. Generally speaking, scanned images, i.e., electronic representations of hard copy documents, are often large, and thus make desirable candidates for compression.

A number of different compression techniques exist, and many of these are proprietary to individual users. The image compression standard disseminated by the Joint Photographic Experts Group (JPEG) committee is a compression technique which reduces data redundancies based on pixel-to-pixel correlations. Generally, picture regions of an image do not change very much on a pixel-to-pixel basis and therefore have what is known as "natural spatial correlation." In natural scenes, correlation is generalized, but not exact. Noise makes each pixel somewhat different from its neighbors.

Lempel-Ziv-Welsh (LZW) encoding, a popular run-length coding standard developed by J. Ziv and A. Lempel, and later refined by T. Welsh, is a technique for compressing data that takes advantage of repetition of strings in the data. Generally, raster data contains a high number of repetition. LZW compression is easy to implement, operates at high speed and results in high compression ratios without loss of data.

## SUMMARY OF THE INVENTION

This invention is directed to the compression of images from documents which contain both pictures and text. Documents containing both pictures and text may be efficiently compressed using the JPEG standard for the picture regions and the LZW encoding technique for text regions.

The disadvantage of hybrid compression involving JPEG standards, for example, is that hybrid compression inherently produces boundary blocks with missing data. JPEG standards typically operate on M×M blocks and the block boundary is generally not aligned with the region boundary. Thus, as shown in FIG. 1, the picture blocks on the boundary, i.e., the boundary blocks **110**, only partially contain picture data **112**, where the rest is either text data or is missing altogether. Since text data cannot be efficiently compressed using the JPEG standard, all non-picture data within these boundary blocks is considered to be "missing data." FIG. 1 shows "missing data" **114** in the boundary blocks **110** occurring in conventional hybrid compression involving the JPEG standard.

The system and method of this invention improve the performance of the hybrid compression process by perform-

ing recursive dilation. This invention provides methods and apparatus that fill in the "missing data" without creating observable artifacts on the decompressed image data.

The compression/decompression method and system of this invention can be implemented, for example, by modifying basic JPEG compression/decompression engines and basic LZW compression/decompression engines. However, it should be appreciated that the compression method and system of this invention can be applied to any set of hybrid compression techniques, where one or both of the compression techniques compresses the image data on a block-by-block basis, such that there are boundary blocks containing more than one type of image data. Accordingly, the compression/decompression method and system can be applied to any number of systems, including digital printers and copiers, that need to provide compressed or decompressed images. A system implementing this invention can include data or image processing systems capable of compressing or decompressing images. The method and system described herein are efficient and result in smoother boundaries between regions of different image data types.

These and other features and advantages of this invention are described in or are apparent from the following detailed description of the preferred embodiments.

## BRIEF DESCRIPTION OF THE DRAWINGS

The preferred embodiments of this invention will be described in detail, with reference to the following figures, wherein:

FIG. 1 shows "missing data" in the boundary blocks occurring in conventional hybrid compression involving JPEG standards;

FIG. 2 is a generalized block diagram of one embodiment of a compression and decompression system according to this invention;

FIG. 3 shows one embodiment of a preferred hybrid compressor according to this invention;

FIG. 4 shows one embodiment of a hybrid decompressor according to this invention;

FIG. 5 is a flowchart outlining one embodiment of an image compression and decompression method according to this invention;

FIG. 6 is a flowchart outlining in greater detail a recursive dilation process according to this invention;

FIG. 7 is a flowchart outlining in greater detail the decompression and pasting process according to this invention;

FIG. 8 shows an example of a recursive dilation process according to this invention;

FIG. 9 shows another example of a recursive dilation process according to this invention;

FIG. 10 shows another example of a recursive dilation process according to this invention; and

FIG. 11 shows another example of a recursive dilation process according to this invention.

## DETAILED DESCRIPTION OF PREFERRED EMBODIMENTS

FIG. 2 shows a generalized functional block diagram of a compression and decompression system **100** according to this invention. The compression and decompression system **100** includes an image source **110** that may be any one of a number of different sources, such as a scanner, digital copier or facsimile device suitable for generating electronic image